

NHẬN DẠNG HÀNH VI MẤT TẬP TRUNG CỦA TÀI XẾ TRONG VIDEO THEO HƯỚNG TIẾP CẬN ĐA GÓC NHÌN

Đào Đức Quang - CH1901027

Tóm tắt



Đào Đức Quang (CH1901027)

- Lớp: CS2205.APR2023
- Link Github:
<https://github.com/quangdilinh/CS2205.APR2023>
- Link YouTube video:

Giới thiệu

- Bài toán:

Nhận diện được thời điểm và phân loại hành vi mất tập trung của tài xế

- Input:

- 3 video về tài xế lái xe ở 3 góc nhìn khác nhau (rear, dashboard, right view).
- Danh sách các loại hành vi cần phân tích.

- Output:

- Thời điểm bắt đầu và kết thúc cùng với nhãn phân loại tất cả hành vi bất thường.

Giới thiệu

INPUT



ID	Description	ID	Description
0	Normal Driving	9	Adjust control panel
1	Drinking	10	Pick up from floor (Dri.)
2	Phone Call(right)	11	Pick up from floor (Pax)
3	Phone Call(left)	12	Talk to Pax (right)
4	Eating	13	Talk to Pax (backseat)
5	Text (right)	14	Yawning
6	Text (left)	15	Hand on head
7	Hair / makeup	16	Singing
8	Reaching behind	17	Shaking or Dancing

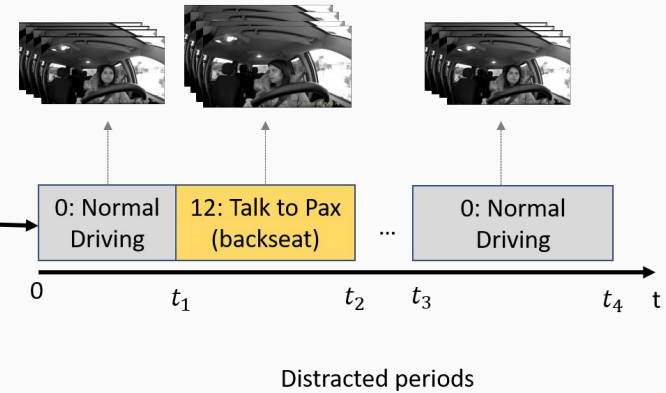
Table 1. 18 distracted behavior actions of SynDD1. Label 0 is not considered for the evaluation.



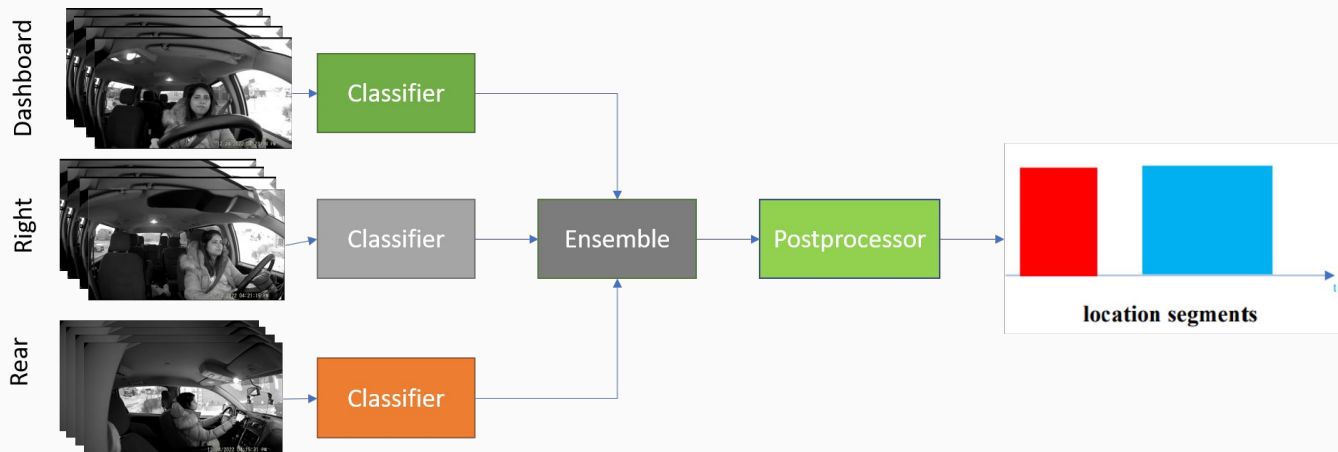
Videos from 3 view

Detector

OUTPUT



Mục tiêu



- Nghiên cứu các thuật toán SOTA và xây dựng giải pháp hiệu quả cho bài toán này dựa vào SOTA với cách tiếp cận dự kiến:
 - post-processing
 - khai thác multi-view.
- Xây dựng được ứng dụng minh họa

Nội dung và Phương pháp

- Cải thiện khả năng nhận diện của mô hình ở các phương pháp đã tìm hiểu
- Đánh giá và so sánh hiệu suất các mô hình bằng độ đo IoU
- Xây dựng chương trình cho phép người dùng input và trả về kết quả nhận diện

Nội dung và Phương pháp

- Khảo sát các paper có hiệu suất cạnh tranh trong việc khai thác multi-view và postprocessing
- Tìm hiểu cấu trúc và cách xử lý của một số thuật toán xử lý đã đạt giải cuộc thi AI City Challenge [3] như:
 - k-fold ensemble [1][4]
 - Nontrivial clustering, custom post-processing [4]
- Thực hiện lại các phương pháp đã tìm hiểu, phân tích kết quả của các giải pháp

Kết quả dự kiến

- Báo cáo các kĩ thuật SOTA đang được sử dụng để tăng cường hiệu quả nhận diện hành vi mất tập trung.
- Kết quả thực nghiệm, so sánh và đánh giá của các phương pháp
- Chương trình minh họa sử dụng mô hình đã được cải tiến

Tài liệu tham khảo

- [1]. Wei Zhou, Yinlong Qian, Zequn Jie, Lin Ma. Multi View Action Recognition for Distracted Driver Behavior Localization. CVPR 2023
- [2]. Zhan Tong, Yibing Song, Jue Wang, and Limin Wang. VideoMAE: Masked autoencoders are data-efficient learners for self-supervised video pre-training. In Advances in Neural Information Processing Systems, 2022.
- [3]. AI City Challenge 2023 <https://www.aicitychallenge.org/2023-challenge-tracks/>
- [4]. Manh Tung Tran, Minh Quan Vu, Ngoc Duong Hoang, and Khac-Hoai Nam Bui. An effective temporal localization method with multi-view 3d action recognition for untrimmed naturalistic driving videos, 2022.